(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 502 011 B1**

(12) # EUROPEAN PATENT SPECIFICATION

(45) Date of publication of patent specification: **08.06.94 Bulletin 94/23**

(51) Int. Cl.^{5}: **C12Q 1/68,** C12N 15/10, // C07K15/16, C07H21/04

(21) Application number: **90917020.1**

(22) Date of filing: **20.11.90**

(86) International application number: **PCT/EP90/02013**

(87) International publication number: **WO 91/07505 30.05.91 Gazette 91/12**

(54) **CLONING METHOD AND KIT.**

(30) Priority: **21.11.89 GB 8926269**

(43) Date of publication of application: **09.09.92 Bulletin 92/37**

(45) Publication of the grant of the patent: **08.06.94 Bulletin 94/23**

(84) Designated Contracting States: **AT BE CH DE DK ES FR GB GR IT LI LU NL SE**

(56) References cited:
EP-A- 200 362
EP-A- 258 017
EP-A- 265 293
EP-A- 297 379
EP-A- 357 011
WO-A-85/04663
WO-A-90/06043
NUCLEIC ACIDS RESEARCH, vol. 18, no. 17, 11 September 1990, Eynsham, Oxford (GB); T. HULTMAN et al., pp. 5107-5112

(56) References cited:
FEBS. LET., vol. 258, no. 1, 20 November 1989, Amsterdam (NL); A.-C. SYVÄNEN et al., pp. 71-74
NUCLEIC ACIDS RESEARCH, vol. 18, no. 20, 25 October 1990, Eynsham, Oxford (GB); C. ASLANIDIS et al., pp. 6069-6074

(73) Proprietor: **DYNAL A.S.**
P.O. Box 158 Skoyen
N-01212 Oslo 2 (NO)

(72) Inventor: **HORNES, Erik**
Lilleakeron 9B
N-0283 Oslo 2 (NO)
Inventor: **UHLEN, Mathias**
Kvarnbogatan 30
S-752 39 Uppsala (SE)

(74) Representative: **Holmes, Michael John et al**
Frank B. Dehn & Co.
European Patent Attorneys
Imperial House
15-19 Kingsway
London, WC2B 6UZ, (GB)

EP 0 502 011 B1

Description

This invention relates to a cloning method and to a kit for performing same.

The ability to splice genes, gene fragments or other target DNA into vectors or other pieces of DNA using restriction enzymes (RE) and ligases has been an important aspect in the advance of molecular biology and biotechnology. However, the technology of recombination or gene splicing has several disadvantages. Firstly, there is the need for conveniently situated restriction sites and often sites have to be constructed which not only takes time but can lead, in the long term, to mismatched reading frames and for example non-expression of a gene of interest. Such sites are usually introduced by means of oligonucleotide linkers which have to be synthesised and purified and are then used in excess to ensure the addition of the required RE sites(s) on to the target DNA. Secondly, the in vitro ligation or splicing of DNA is rather inefficient and relatively cumbersome screening techniques are required to locate desired recombinants. Finally, the technology is time-consuming and is not well-suited to automation. Accordingly, there is a need for a simple and relatively rapid method of cloning which avoids the problems of conventional splicing and the use of conventional plasmids.

It should be noted that conventional plasmids for cloning normally take the form of double stranded cyclic plasmid structures containing a promoter region separated from a gene or other DNA sequence of interest for replication or expression by one or more RE sites which permit the DNA of interest to be excised subsequently; such sites are also used for insertion of the DNA of interest for replication or expression, via one or more (RE) sites in the linearised plasmid which permit the introduction of DNA of interest, which is provided with 'sticky ends' corresponding to RE sites of the linearised plasmid. When DNA has been synthesised, for example by cDNA synthesis from mRNA, by mutagenesis or by chemical synthesis, it is in single stranded form which is then conventionally treated with a polymerase to synthesise the second strand, provided with the required 'sticky ends', inserted into the double stranded plasmid vector and ligated to join covalently the insert to the vector which is then used to transform a host microorganism, e.g. E. Coli.

In recent years, the polymerase chain reaction (PCR) has been used for the amplification of target DNA. While this produces increased amounts of the DNA, it is often required to produce larger quantities by cloning in a suitable vector using a host microorganism such as E. Coli. Furthermore, for production of the corresponding protein it is required to incorporate the DNA into an expression vector.

For the reasons given above, conventional techniques for splicing the target gene into plasmid vectors are time consuming and inefficient and not well suited to automation. In cases where PCR itself is effected by an automated technique, it would be desirable for incorporation into the vector also to be readily added on to the automated system.

The present invention has as an object a method which provides for the formation of recombinant DNA from PCR amplified DNA without the need for restriction enzymes or ligases or the provision of restriction sites.

Accordingly, the present invention provides, in one aspect thereof, a method of amplifying target DNA wherein said DNA is first amplified by PCR, the amplified DNA then being contacted with a single stranded linearised plasmid vector having terminal regions which are complementary to terminal regions of the PCR amplified DNA, whereby a cyclic product is formed comprising single stranded sequences from said target DNA and said vector and two double stranded regions from the overlapping terminal regions of the vector and the PCR-amplified DNA; the cyclic product then being introduced into a host organism.

It is surprising that the cyclic product can be used to transform a host directly; the native enzyme system of the host organism is capable of chain extension to complete synthesis of the double stranded plasmid which is then available for replication and/or expression of the DNA of interest.

An advantage of the method according to the invention over conventional PCR is that the target DNA is first amplified by PCR sufficiently to give enough DNA for practical purposes of transformation of a host. The host cell replicates the target DNA quite rapidly but highly conservatively, and without using expensive chemicals such as nucleoside triphosphates. The conservation in amplification is important since conventional PCR is known to suffer from errors introduced by mis-matched codons. Not only are such errors amplified during each cycle of PCR but more errors are created in each cycle and this creates a high background level of contamination. Cloning in a host organism can be used to detect errors and select only accurately amplified DNA.

The complementary regions of the PCR amplified DNA may be present in the target DNA but advantageously they are provided as single-stranded nucleotide extensions on the PCR primers, which extensions or 'handles' do not bind to the target DNA (as described in co-pending International patent application WO 90/11369.

Preferably the PCR amplification step of the method according to the invention involves nested primers, as described in the above co-pending PCT application, and this leads to greater sensitivity in isolating and amplifying the target DNA.

Advantageously, the terminal overlapping regions are sufficiently large to provide an adequate hy-

bridisation overlap between the PCR amplified DNA and the terminal regions of the linearised single stranded plasmid so as to form a stable cyclic product, yet still reasonably short in order to avoid unnecessary chemical synthesis, if using primer extensions. It will be clear to persons skilled in the art that the size and stability of the overlapping regions will be dependent to some degree upon the ratio of A-T to C-G base pairings since more hydrogen bonding is available in a C-G pairing. Also, it will be apparent that the smaller the overlap the more likely there will complementarity with a non-terminal region and that if the terminal regions get too large there is always the possibility that the single strand will fold back on itself and hybridise to give a hairpin or dumbbell structure. It is preferred that the overlap should be at least a ten base pair overlap, more preferably at least a twenty base pair overlap.

The term "cloning vector" as used herein includes plasmid vectors both for simple replication and for expression. A replication vector will contain an origin of replication and usually a marker e.g. an antibiotic resistance marker, to aid recognition. An expression vector will normally contain promoter and initiator sequences which must be operably connected in the same reading frame as the DNA insert if this is to be expressed correctly, as well as operator and expression control sequences and a ribosome binding site; appropriate markers e.g. antibiotic resistance markers, are usefully present. In both cases, appropriate RE sites for excision of the DNA will be desirable, especially in the replication vector.

As indicated above, the cyclic product, which is essentially single stranded apart from the two overlapping regions, may surprisingly be used after annealing to transform a host microorganism, thus avoiding the steps of second strand synthesis, and modification and ligation. Since such DNA is synthesised chemically in single stranded form, the simplified procedure of the invention lends itself to the rapid cloning of DNA so synthesised. Also, the method according to the invention is well suited to automation since no steps of precipitation, extraction, filtration, centrifugation or treatment with enzymes are required in getting the PCR-amplified DNA into a host cell.

The linearised single stranded vector may conveniently be a standard vector the terminal sequences of which comprise one or more RE sites permitting a variety of restriction endonucleases to be used to excise the DNA of interest after replication.

Most cloning vectors now in use have a common ancestry, e.g. pUC18, and include the so-called multiple cloning site including several RE sites flanked by longer regions which are also identical. In the case of pUC18, the flanking regions are part of the E. coli Lac Z gene. It may be convenient to include the multiple cloning site or at least one RE thereof with the PCR amplified DNA insert and to use the two flanking regions as the overlapping sequences in accordance with the invention. It is thus convenient to provide the DNA insert with terminal regions complementary to such standard sequences. the term "complementary" as used above means that the regions hybridise in the correct orientation to form the required cyclic product in which the overlapping 3' ends can serve as primers for synthesis of the remainder of the complete double stranded vector by the host organism.

The PCR-amplified target DNA can be cloned into different vectors provided that complementary overlap regions exist between the vector and the DNA. This is significant, for example, where it is desired to insert a gene fragment into many different vectors, such as expression vectors.

In general, to ensure adequate hybridisation and stability of the cyclic product, the overlapping regions are preferably 20 to 250 base pairs (bp) in length or even longer (e.g. 500 bp), more preferably 40 to 200 bp. However, if the overlapping regions are too long, the length of the region to be amplified may be limited in view of the fact that PCR is most effective in the region of 500 to 2,000 bp.

The hybridisation reaction is preferably effected in a 1M sodium chloride solution or an equivalent solution known in the art. (Nucleic Acid Hybridisation, B D Hames and S J Higgins, IRL Press, 1985).

In the PCR stage, the unamplified target dsDNA is first denatured and primers are annealed to both the coding and the non-coding strand. The primers are preferably those corresponding to the 5'-terminal sequences of the DNA so that on extension of the primer with a polymerase, the whole target DNA sequence of each strand will be replicated. The double stranded DNA so produced is then denatured by raising the temperature followed by rapid cooling. An excess of the primer molecules is present and these are annealed to the newly formed coding and non-coding strands. Extension using polymerase produces further double stranded DNA. The temperature cycling can be repeated many times, thereby producing a large number of copies of the DNA. Preferably, the polymerase used is one which can withstand the highest temperature of the cycle, commonly the Taq polymerase, otherwise there is a need to separate the polymerase from the nucleic acids before each heating step or replenish the polymerase after each cooling step. It is also preferred that the polymerase has a high proof-reading ability to avoid mis-matched bases and randomly introduced errors. An example of such a polymerase is vent polymerase available from New England Biolabs. Such PCR amplification provides target DNA incorporating the primers which are used. As mentioned above, nested primers may be used. In this case PCR is carried out with a first set of primers for a given number of cycles e.g. about 25. The amplified DNA is then contacted with a sec-

ond pair of primers, one or both being different from the primers used earlier and being inboard of the binding sites of the first primers.

Since the method of the invention uses single stranded amplified DNA, it is advantageous for one PCR primer to carry means for immobilisation, e.g. a biotin molecule, or to be already attached to a support.

The double stranded amplified DNA may then be subjected to strand separation whereby one strand (unwanted) remains immobilised while the other is liberated into solution and can be combined with the linearised vector in accordance with the invention. Thus, such strand separation after PCR is an important preferred aspect of the invention.

However, since the linearised standard vector will hybridise only to one of the two PCR amplified DNA strands, it is also possible to liberate both strands into solution by conventional strand separation and to react these directly with the linearised standard vector. This will, however, be less efficient due to competing re-assembly of double stranded target DNA.

The PCR stage of the invention may also include a subsequent step of site-specific mutagenesis of the target DNA. In one such strategy, a standard linearised single stranded vector can be prepared by taking a plasmid in double stranded form containing two outer RE sites and two further inner RE sites inside these, each separated from the outer RE sites. The plasmid is cut at one of the inner RE sites and biotinylated followed by restriction at the other inner RE site. This provides a linearised double stranded vector which is then attached to an insoluble support coated with avidin or streptavidin. One strand of the linearised vector is thus attached to the support which the other is not and the latter can then be brought into solution by denaturation. In this example, a further plasmid contains the DNA sequence to be mutagenised flanked by sequences corresponding to the terminal regions of the single stranded linearised vector. For example the further plasmid may be the standard vector having the DNA sequence to be mutagenised inserted between the two inner RE sites. The further plasmid is subjected to at least one or two cycles of PCR amplification using primers flanking the target DNA sequence (to be mutagenised), these primers being homologous with the terminal sequences of the linearised vector. For example, the primers may correspond to the sequences between the outer and inner RE sites of the standard vector. One of the primers is provided with means for attachment to a support (e.g. a biotin group) or is already attached to the support. Chain extension provides, after a final strand separation, the target DNA in single stranded form linked at one end to a support. Hybridisation to a further primer at the 3′-end to initiate chain extension and a mutagenesis primer incorporating the desired mutation, permits synthesis, in the presence of a polymerase, of a DNA strand incorporating the mutation and flanked by terminal sequences complementary to those of the linearised vector. Strand separation, e.g. by treatment with alkali, liberates the mutagenised strand into solution while the template is immobilised and thus readily separated. The mutagenised ss DNA may then be contacted with the linearised vector and annealed to give a cyclic product in accordance with the invention.

The target DNA may be cDNA produced by reverse transcription from mRNA, and the method according to the invention therefore provides a way of direct cloning cDNA. For example, mRNA may be immobilized on a solid support bearing poly dT which hybridises to the poly A tails of the mRNA.

Preferably, the connection of the poly dT to the solid support includes a suitable RE site. Reverse transcription can then be effected advantageously using the poly dT as a primer. The mRNA is then removed leaving the newly synthesised single-stranded cDNA attached to the solid support.

The single-stranded cDNA may be made double-stranded by use of a suitable polymerase, e.g. T4 polymerase, and the free end of the cDNA may have attached thereto a linker using a suitable ligase. In this case the linker and the sequence proximate the solid support are, advantageously, complementary to the terminal regions of the plasmid vector. The double-stranded cDNA can then be subjected to PCR in accordance with the invention; the PCR primers corresponding to the linker sequence and the sequence proximate the solid support. Alternatively, the linker and/or sequence proximate the solid support may not be complementary with the terminal regions of the vector in which case such complementary regions can be provided by using PCR primers with handles.

Instead of forming double-stranded cDNA as mentioned above, it is possible to use a terminal transferase to add several molecules of one type of nucleotide to the 3′ end of the single-stranded cDNA to form a tail, for example a dG tail. Thus the single-stranded cDNA comprises a 5′ poly dT sequence near the support (which sequence hybridised to the poly A tail of the mRNA) and a 3′ tail, for example poly dG. PCR can now be initiated using poly dT and poly dC primers. The linearised single-stranded vector preferably has complementary terminal poly dA and poly dC regions in order to form the cyclic product with the PCR amplified single-stranded cDNA. As mentioned above, it is of course possible to use primers in the PCR amplification which comprise either poly dT or poly dC and appropriate handles to form the overlap with complementary regions of the vector.

The method according to the invention in combination with direct solid phase DNA sequencing can be used to analyse target DNA, for example alleles of a

locus e.g. the human apolipoprotein E locus and may thus be used for diagnosis of physiological conditions. Such analysis may include sequencing, as will be exemplified below. Moreover, the method according to the invention can be used in a diagnostic method, for example testing for the presence of a certain allele. For example, the cloning may be followed by direct sequencing to separate and identify both alleles in a heterozygote. Such direct clinical sequencing to detect polymorphism has the advantages that non-expected nucleotide changes in close proximity to the allele analysed will be detected and that flanking sequences can be used as positive controls in order to verify that the non-expected exchange was not due to the earlier PCR amplification.

Clearly, in such methods as above where one is investigating genomic DNA, overlapping terminal regions are provided by the primers during amplification; there is no need to provide specific terminal RE sites to allow incorporation of the target DNA into a vector (as required in conventional cloning and amplification protocols). However, it is preferably that RE sites are incorporated adjacent the target DNA since such RE sites will allow for subsequent excision of the cloned target DNA from the vector. RE sites can be conveniently provided in the overlapping terminal regions provided by the primers during amplification. If, for example, the vector includes a multiple cloning site having many RE sites, one site can be selected for restriction and the overlapping regions of the primers will then be complementary to the terminal regions either side of the selected restricted site.

It should be noted that it is preferable to remove any excess primer remaining after PCR amplification since otherwise it will compete with the terminal regions of the plasmid to hybridise with the amplified DNA.

It will be appreciated that in any of the above systems, the biotin/avidin or streptavin affinity coupling may be replaced by other such coupling using a relatively small molecule and binding partner, e.g. an antigen and antibody therefor, or covalent coupling as indicated below.

An advantage in PCR strategies involving immobilised site-specific mutagenesis is that the template is readily removed completely from the synthesised DNA, thus avoiding contamination with unmutated DNA.

As mentioned earlier, it may be convenient in some of the above PCR strategies to use PCR primers having 'handles' of DNA not hybridising in the first cycle of the PCR amplification, such handles corresponding to the terminal regions of the standard vector while the hybridising regions of the primers correspond to regions of a source of target DNA e.g. genomic DNA. This applies equally to PCR amplification with or without mutagenesis.

The insoluble support, where used, may take a Variety of forms, for example microtitre wells, filters made from materials such as cellulose or nylon, or particles including, for example, sephadex or sepharose beads or polystyrene latex particles. It is a preferred feature of the invention to use magnetic particles which may be aggregated onto a surface and then be readily redispersed for a subsequent treatment step, e.g. by physical agitation.

Probe and primer oligonucleotides may be prepared by using any of the commercially available DNA synthesis devices, e.g. those available from Applied Biosystems, Inc. (850-T Lincoln Center Drive, Foster City, CA 94404).

Some aspects of the process of the invention are in part disclosed in our International Patent Application No.W090/06043.

The invention also includes kits for carrying out the cloning procedure of the invention comprising one or more of the following:-

a) a standard linearised vector in single stranded or double stranded form the said double stranded form immobilised by one end of one strand thereof.

b) an insoluble support carrying one member of a pair of binding partners.

c) nucleotides carrying the other member of said pair of binding partners.

d) a polymerase.

e) 2 PCR primers corresponding to the terminal regions of said standard vector one of which is adapted to bind to said support.

f) a thermostable polymerase.

g) an alkaline solution for strand separation.

In particular, the invention provides a kit for amplifying target DNA comprising:-

a. a linearised vector in single stranded or double stranded form or said double stranded vector in linearised form immobilised by one end of one strand thereof;

b. a polymerase;

c. two PCR primers corresponding to the terminal regions of said vector; and

d. triphosphate nucleosides.

Preferably, such a kit additionally comprises means for sequencing target DNA including either labelled primer or labelled nucleoside triphosphates. Also, such a kit may additionally comprise means for site-specific mutagenesis including:-

i. a site-specific mutagenic primer;

ii. a second polymerase; and

iii. a ligase.

The invention will now be described by way of non-limiting examples with reference to the drawings in which:-

Fig. 1 shows a protocol for site-specific mutagenesis using the method according to the invention;

Fig. 2 shows a protocol for amplification of genomic DNA using the method according to the in-

vention;

Fig. 3 shows a region of a human lipoprotein E gene together with primers; and

Fig. 4 shows sequencing printouts for two clones.

The terms Sepharose, Pandex, Aminolink, Tween and Autoread are registered trademarks in at least one contracting state.

Example 1

In vitro mutagenesis on latex particles.

The protocol shown in Figure 1 was used

(a) To yield the ss vector template 10 μg of pUC18 were digested with EcoRI in a total volume of 50 μl. The 5′ protruding ends were filled in using Klenow polymerase (5 U) and 2 μl Biotin-7-dATP (BRL), 7.5 μl of a buffer containing 100mM Tris-HC1 (pH 7.5), 100 mM $MgC1_2$ and 1M NaC1. The volume was adjusted to 75 μl with water. The reaction was performed at room temperature during one hour and after that purified using a Sephadex® G50 spin column. The purified biotinylated linearized vector was cut with HindIII. The reaction containing the biotinylated double stranded DNA was mixed with previously washed Pandex avidin particles, (Baxter Healthcare Corp., Mundelin, Illinois, USA).

To yield the ss vector template the immobilized doublestranded DNA was converted into singlestranded form by melting off the non attached strand by incubation at 37°C with 20 μl 0.15 M NaOH for 10 minutes. The pH of the supernatant was immediately adjusted with 1.5 μl HAc (1.7M) and 2.2 μl 10 x TE (100mM tris pH 7.5, 10 mM EDTA).

(b) To yield the mutagenesis template, the inserted fragment from pUCRA was PCR amplified using 10 pmol of primer A TGC-TTC-CGG-CTC-GTA-TGT-TGT-GTG3′ and biotinylated primer B Biotin-AAA-GGG-GGA-TGT-GCT-GCA-AGG-C GA3′ in 25 μl PCR reaction mixture as recommended by Perkin Elmer and amplified for 20 cycles. After PCR amplification, Pandex avidin particles were added to immobilize the amplified insert with flanking vector sequences.

(c) To yield template for in vitro mutagenesis the immobilized PCR amplified fragment was made single stranded with 0.15M NaOH for 10 minutes at room temperature. 10 pmol were added of each primer Q 5′CGG-CTC-GTA-TGT-GTG-GAA-TTG and mutagenesis primer M 5′CC-AAT-GCA-TAT-GTG-GTC-GGC-TAC-GCT-GGA-AA T-AGC-GCA-TAT-TTC3′ (the orginal sequence was: CCAAT GCA-TAT-GTG-GTC-GGC-TAC CGT GCT GGA AAT AGC-GCA) were annealed to the template immobilized on the Pandex® avidin particles in a solution containing 10 mM Tris-HC1 (pH- 7.5), 10 mM $MgC1_2$, 100 μg/ml BSA and 100 mM NaC1. The mixture was incubated at 65°C for a few minutes and allowed to cool to room temperature.

(d) Extension was performed by adding 1 μl BSA (100 μg/ml), 6μl polymerase mix (100 mM Tris-HC1 pH 8.8, 10 mM DDT), 50 mM $MgC1_2$ and 5 mM ATP), 6 μl chase (10 Mm each of dNTP) and 3.5 U T_4 DNA polymerase. 1 unit of T_4DNA ligase was added to the beads containing the insert strand.

The volume was adjusted to 30 μl with water. The mixture was incubated at RT for 20 minutes followed by incubation on a roller mixer at 37° during two hours.

(e) After extension the beads were washed once with TE. The newly synthesized strands were melted off by incubation with 20 μl 0.15M NaOH at 37° during 10 minutes. The pH of the supernatant was immediately adjusted with 1.5 μl HAc (1.7M) and 2.2 μl 2 x TE.

The two supernatants, the single stranded vector and the newly mutated insert with flanking vector sequences were mixed and incubated at 70°C for 10 minutes and allowed to cool to RT.

After annealing of the two strands the $CaC1_2$ concentration was adjusted to 0.1M and E. coli DH5α was transformed with DNA and spread on TRAB plates containing IPTG and X-Gal.

Example 2

Direct cloning of the human genomic apolipoprotein E gene using magnetic separation of single stranded DNA.

Materials and Methods

Clinical samples

Leukocytic DNA from venous blood samples from patients having the genotype E2/4 was kindly provided by A.-C. Syvänen and K. Aalta Setälä (1) (Orion Corporation, Helsinki, Finland).

Preparation of the primers

Four PCR primers (denoted RIT113, 114, 123 and 125) were synthesized on an Applied Biosystems 381A DNA synthesizer. 5′-amino modified oligunucleotides (RIT123) were synthesized using the reagent Aminolink® 2 (ABI, USA). A biotin residue was attached to the amino group using the reagent Biotin-X-NHS ester as described by the manufacturer (Clontec, USA) and the biotinylated oligonucleotide was purified by reverse phase HPLC. The fluorescent M13 universal sequencing primer was purchased

from Pharmacia LKB Biotechnology, Sweden. The two primers RIT 123 and RIT 125 contained a 5'sequence (22 nucleotides) complementary to the pUC vector.

Polymerase chain reaction

The DNA (100 ng per sample) was amplified with the RIT113 and RIT114 primers at final concentration of 1 μM. The PCR was carried out in 100 μl of a solution consisting of 0.2mM each of the dNTP:s 20mM Tris-HC1, pH 8.8, 15mM $(NH)_2SO_4$ 1.5mM $MgC1_2$ 0.1% Tween® 20,0.1 mg/ml gelatin and 2.5 units of Taq polymerase (United States Biochemical Corp, USA) in a thermal cycler (Perkin-Elmer, USA) for 25 cycles of 1 min. at 96°C and 2 min. at 65°C. For amplification with a second pair or primers a small aliquot (3 μl of 1:100 μl dilution) of the PCR product amplified with the primers RIT113 and RIT114 was transferred to a second PCR. This was carried out at the conditions described above using the biotinylated primer RIT123 and the primer RIT125 at 0.1 μM concentration.

Immobilization on magnetic beads

Magnetic beads containing covalently attached Streptavidin. Dynabeads ® M280 Streptavidin, were obtained from Dynal (N-0212 Oslo 2, Norway). A neodynium-ironboron permanent magnet MPC-E, (Dynal, Norway) was used to sediment the beads in the tubes during supernatant removal and washing procedures. The PCR mixture was added to 300 μg of Dynabeads ® M280 Streptavidin previously washed with TE buffer (10mM Tris pH 7.5, 1 mM EDTA) containing 1M NaC1, and incubated 15 min. at room temperature.

Preparation of single stranded vector

A total of 5 μg of pUC18 (Pharmacia LKB Biotechnology, Sweden) was digested with EcoRI, phenol extracted followed by desalting with a sephadex G50 spin column. Biotin-7-dATP (BRL, USA) was introduced by Klenow polymerase. After heat inactivating and desalting it was digested with HindIII and immobilized on 1 mg of Dynabeads ® M 280 Streptavidin. Strands were separated using 40 μl 0.12 M NaOH. The supernatant was neutralized by adding 3.7 μl 1.7 M HAc and 4.4 μl 10 x TE buffer pH 7.5. The concentration was estimated using agarose gel electrohoresis.

Direct cloning

The PCR amplified apoE gene region containing ends complementary to the single stranded vector was immobilized on magnetic beads. The strands were separated using 40 μl 0.12 M NaOH. The supernatant was neutralized by adding 3.7 μl 1.7 M HAc and 4.4 μl 10 x TE buffer PH 7.5 and concentration estimated. 100 ng single stranded vector was mixed with an equal amount of single stranded genomic amplified DNA in a total of 5 μl. Transformation of competent E. coli DH5α cells (BRL.USA) was performed according to the manufacturer's direction.

Sequencing reactions

The immobilized dsDNA was washed with 50 μl TE buffer and then incubated with 10 μl 0.1 M NaOH for 15 min. at room temperature. The supernatant was removed and the beads containing the immobilized single stranded DNA were washed with 50 μl 0.15 M NaOH and 3 times with 50 μl TE buffer. The volume was adjusted to 13 μl with H_2O .

All sequencing reactions were performed with reagents from the AutoRead® sequencing kit (Pharmacia LKB Biotechnology, Sweden). 2 μl(1 pmol) of a fluorescent labelled universal sequencing primer was added to each Eppendorf tube together with 2 μl of annealing buffer. The annealing mixtures were incubated at 65°C for 15 min. and allowed to cool to room temprature for 15 min. 1 μl of a MID solution (Manganese, Isocitrate and DTT) was added to each annealing mixture together with 2 μl T7 polymerase (1.5 units/μl) and 2.5 μl of respectively A,C,G,T sequencing mixture (containing c^7dGTP instead of dGTP) were prewarmed at 37°C for 1 min. using a microtest plate (Sarstedt, West Germany) before 4.5 μl of the annealing mixture was added to each sequencing mixture and incubated for 5 min. at 37°C. 5μl of deionized formamide containing Blue Dextran was added to stop the reactions. the microtest plate was heated at 80°C for 2 min. and 5 μl was loaded on a 6% sequencing gel run on an automatic sequencing apparatus with detection of fluorescent bands during electrophoresis (A.L.F, Pharmacia LKB Biotechnology, Sweden).

Results

The principle for solid phase cloning

A basic concept for cloning using a solid phase approach is shown in figure 2. A single stranded vector fragment is provided by selectively incorporating biotin into one of the strands of the vector DNA. This is achieved by restriction and fill-in using a biotin-dNTP and DNA polymerase. The double stranded DNA is bound to magnetic beads containing streptavidin and the single stranded vector is simply eluted with alkali. If more single stranded vector fragment is needed, a run-off extension reaction with DNA polymerase can be carried out, one or several times, and the extended fragment is again eluted with alkali and

collected. This yields a well defined single stranded vector with flanking sequences represented by the A and B region (fig.2).

Alternatively, the vector fragment can also be prepared by the apparently simple PCR procedure using specific vector primers. However, caution must be taken as PCR of large fragments might create random polymerase errors into the vector part which is not easy to control. In addition, PCR of larger fragments (> 3 Kb) is not yet straight forward in terms of reproducibility and yield.

The "insert" DNA to be cloned is obtained by PCR, using specific primers with handles consisting of the vector regions A and B, respectively (Fig.2). For genomic DNA specific primers need to be synthesized, while for gene fragments inserted into vectors it is possible to use general PCR primers designed for the solid phase cloning. In both cases, one of the primers contains a biotin in the 5′ end, which allows the in vitro amplified material to be captured by Streptavidin coated magnetic beads. A single stranded insert fragment with flanking regions (A′ and B′) complementary to the vector fragment can subsequently be eluted th alkali. The two single stranded fragments can then be mixed to form a gap-duplex molecule (fig.2) and transformed directly in E. coli. The method allows cloning of any fragment, from any origin (chromosomal DNA, plasmid DNA etc.) independent on restriction sites. No restriction enzymes and ligase are needed and very high yield of recombinants is expected since the vector and insert alone should not give transformants.

Design of primers for the human apolipoprotein E gene

The solid phase approach was tested by analysing the chromosomal gene fragments of the apolipoprotein E (ApoE) gene. Mature ApoE is a 299 amino acid protein which plays an important role in the lipoprotein metabolism (2). In humans, three major apoE isoproteins exist (3), apoE2, E3 and E4, encoded by the three different co-dominant alleles (E2, E3 and E4). Besides apoE2 (cys_{112}), cys $_{158}$) E3 (cys_{112}, arg_{158}) and E4 (arg_{112}, arg_{158}), several rare, independent apoE isoproteins have recently been described in this region (4). Sequencing of individual chromosomes is therefore of importsance to establish the structure of the alleles in this region on both chromosomes.

The polymorphic region of apoE is due to single base substitutions at two loci, a C/T nucleotide change at codon 112 and a similar C/T nucleotide change at codon 158. Both mutations give rise to an arginine to cysteine replacement (fig.3). To test the solid phase cloning protocol (fig.2), a nested primer approach was followed with a pair of outer primers and another pair of inner primers (fig.3) used for the cloning (RIT 123 and RIT 125) and thus containing handle sequences complementary to the pUC vector. The downstream primer (RIT123) contains a 5′-biotin to allow capture of the amplified chromosomal DNA.

Direct solid phase cloning of the chromosomal apoE gene

Blood samples from several human patients were prepared (1) and used for a two step (25 cycles each) PCR procedure using the outer primers (RIT 113 and RIT 114) and the inner primers (RIT123 and RIT 125). For details see Materials and Methods. Analysis by agarose electrophoresis showed a band of the expected 290bp for all samples (data not shown).

The PCR product for one of the patients was bound to the magnetic beads and the single stranded insert was eluted with alkali and neutralized. The vector fragment from pUC18 was prepared by restriction with EcoRI, followed by a fill-in reaction with biotin-dATP and by restriction with HindIII. After binding to magnetic beads, the single stranded vector was eluted with alkali and neutralized.

The single stranded vector and the eluted single stranded PCR fragment were mixed and used directly to transform competent E.coli cells. Several hundred colonies were obtained, while transformation with the vector or the insert alone, gave a few or none colonies, respectively (data not shown). Restriction mapping of purified plasmids from 20 colonies showed that 19 recominants had plasmid with an insert of the correct size (data not shown).

Sequencing of positive clones

Six of the colonies from the mixing experiment were sequenced directly by the solid phase method (5). Examples of two of the samples are shown in figure 4. The sequence data show that one of the clones (Fig.4A) has a G/G in the polymorphic positions (arrows) corresponding to a C/C loci. In contrast the other clone has an A/A (Fig.4B) in these positions, suggesting a T/T genotype. Of the six clones sequences, four were of the T-T genotype, while two were of the C-C type. Clearly, the patient is a heterozygote E2/4 with an arg_{112}, arg_{158} coded by one of the chromosomes, while the other chromosome codes for an apoE protein with a cys_{112}, cys_{158}. An interesting observation is that four of the six clones sequenced showed different unique sequences outside the allelic codons, such as the T at position 127 in figure 4B. These nucleotide changes correspond to polymerase errors obtained during the repetitive PCR procedure. Two of the clones, such as the one showed in figure 4A,contained a sequence without any random errors.

DISCUSSION

We have for the first time shown that magnetic separation of DNA can be used for efficient assembly of recombinant DNA molecules. Here, individual human chromosomal gene fragments were directly cloned by PCR simply by using a 22 basepair primer "handle" complementary to the ends of the linear vector fragment during the PCR. The cloning step is simple, rapid and involves no ligation and restriction enzyme reactions. the yield of specific chromosomal clones was approximately 90% without using any positive selection or any special host strain.

The fragment produced can be cloned into different vectors provided that complementary overlap regions exist between the vector and the insert. This allows for the use of a battery of prepared single stranded vectors, in which the insert is directly cloned simply by mixing and transforming. This procedure is well suited for automation since no precipitations, extractions, filtrations or centrifugations are needed and no enzymatic steps are performed. Thus both manual, semi-automated procedures can be envisioned. This is significant for large scale projects, where it is desired to insert a gene fragment into many different vectors, such as various expression vectors.

The cloning protocol (fig.2.) has the advantage that the same principal result can be accomplished both with or without using PCR to prepare the linearised vector. This is important since accumulated polymerase errors are a major concern when ever PCR products are cloned (6), which makes it strongly desired to sequence the cloned material. For large sized fragments such as cloning and expression vectors, this is difficult and time consuming. Therefore, solid phase cloning protocols that do not depend on PCR produced vector are attractive. In this Example, the vector was produced by the restriction-fill-in procedure to avoid PCR amplification of larger sized fragments while the cloned chromosomal ApoE gene was obtained by PCR.

As expected, the cloned material has considerable amounts of randomly introduced errors (Fig.4.). As the error frequency of Taq polymerase is approximately 10^{-4}(2) and the PCR was carried out for 2x25 cycles, the theoretical error frequency for each nucleotide is roughly 1 out of 200 (10.000/50). This background (less than 1 per cent) is obviously not observed when a direct genomic sequencing is carried out. Even if only one template molecule is present in the original sample and an error is introduced in the first cycle, the correct signal at that position is theoretically 75% and can possibly be discriminated from background.

In contrast, when the same material is cloned, for example in E. coli, the random errors become prominent readily detectable. For a fragment of the size of 200 basepairs (as the human apoE gene described here) and error frequency of 1/200 means that most fragments will contain an introduced error. The results of the cloning confirm this as 4 out of 6 cloned fragments contained random errors. The high frequency of introduced errors detected in the fragments after cloning can of course be limited by performing less PCR cycles or to avoid the nested primer approach. It might also be possible to use a less error prone polymerase. However, as long as relatively short fragments are cloned and a straight forward sequencing of several clones can be performed, it should be possible to find a clone with a correct sequence by a small scale screening. Note, that the correct genomic sequence can be determined and defined by the direct genomic solid phase sequencing (5).

Interestingly, the cloning procedure followed by direct sequencing using an automated electrophoresis instrument can be used to separate and identify both alleles in a hetrozygote. Thus, a diagnostic evaluation may be performed in an automated manner, which is in contrast to most polymorphic analysis based on hybridization (7). A direct clinical sequencing approach for diagnosis of polymorphism has the additional advantages that non-expected nucleotide changes in close proximity to the allele analyzed will be discovered and that the flanking sequences can be used as positive control to show that the PCR reaction has been successful and specific.

In conclusion, the direct cloning procedure according to the invention was able to isolate and sequence individual human chromosomal apoE gene fragments. Thus, the relationship between two separated alleles could be resolved. The results demonstrate the selectivity and efficiency obtained by the solid phase approach as all the recombinants sequenced had the desired chromosomal gene fragment. The cloning using magnetic separation is thus a highly efficient, rapid and simple tool to obtain recombinant molecules, although caution must be taken to minimize the effect of random errors introduced during the PCR by the Taq polymerase. This method and similar procedures can facilitate considerably the assembly of cloned genes in molecular biology and biotechnology.

References:

1. Syvänen, A.-C., K. Aalto-Setälä, K. Kontula and H. Söderlund. 1989. Direct sequencing of affinity-captured human DNA:application to the detection of apolipoprotein E polymorphism. FEBS Lett. 258:71-74.

2. Mahley, R.W. 1988 Apolipoprotein E: Cholesterol Transport Protein with Expanding Role in Cell Biology. Science 240:622-630.

3. Zannis, V.I., P.W. Just and J.L. Breslow. 1981. Human apoliprotein E isoprotein subclasses are

genetically determined. Am J. Hum. Gent 65 232-236.
4. Paik, Y.-K., D.J. Chang, C.A. Reardon, G.E. Davies, R.W. Mahley and J.M. Taylor, 1985. Nucleotide sequence and structure of the human apolipoprotein E gene. Proc. Natl. Acad. Sci. USA 82:3345-3449.
5. Hultman, T., S. Stahl, E. Hornes and M. Uhlen. 1989. Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads and solid support. Nucleic Acids Res. 17:4937-4946.
6. Tindall, K.R. and T.A. Kunkel. 1988 Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase, Biochemistry 27:6008-6013.
7. Caskey, C.T. 1987, Disease Diagnosis by Recominant DNA Methods. Science 236:1223-1229.

Claims

1. A method of amplifying target DNA wherein said DNA is first amplified by PCR, the amplified DNA then being contacted with a single stranded linearised plasmid vector having terminal regions which are complementary to terminal regions of the PCR amplified DNA, whereby a cyclic product is formed comprising single stranded sequences from said target DNA and said vector and two double stranded regions from the overlapping terminal regions of the vector and the PCR amplified DNA; the cyclic product then being introduced into a host organism.

2. A method of amplifying target DNA as claimed in claim 1, wherein the PCR is two-stage and uses nested primers, the or each primer of the second stage being complementary to a sequence of the target DNA between the sites complementary to the first stage primers.

3. A method as claimed in claim 1 or in claim 2 wherein the or the second stage primers comprise single-stranded nucleotide extensions which are complementary to the terminal regions of the vector.

4. A method as claimed in any one of the preceding claims, wherein each of the overlapping region of the cyclic product comprises at least one restriction site.

5. A method as claimed in any one of the preceding claims wherein one PCR primer carries means for immobilisation or is already immobilised.

6. A method as claimed in any one of the preceding claims wherein site-specific mutagenesis is effected between PCR amplication and formation of the cyclic product.

7. A method as claimed in claim 6, wherein PCR amplified target DNA is immobilised prior to mutagenesis.

8. A method as claimed in any one of the preceding claims including the subsequent step of sequencing the target DNA after cloning.

9. A method as claimed in any one of the preceding claims wherein the single stranded linearised plasmid vector is produced by end-immobilising one strand of a double stranded linearised vector, followed by a strand separation step.

10. A method of diagnosis wherein said target DNA is characteristic of a physiological condition and the amplified and cloned DNA produced by the method of claim 1 is subsequently characterised.

11. A kit for amplifying target DNA comprising:-
    a. a linearised vector in single stranded or double stranded form or said double stranded vector in linearised form immobilised by one end of one strand thereof;
    b. a polymerase;
    c. two PCR primers corresponding to the terminal regions of said vector; and
    d. triphosphate nucleosides.

12. A kit as claimed in claim 11 additionally comprising means for sequencing target DNA including either labelled primer or labelled nucleoside triphosphates.

13. A kit as claimed in claim 11 additionally comprising means for site-specific mutagenesis including:-
    i. a site-specific mutagenic primer;
    ii. a second polymerase; and
    iii. a ligase.

Patentansprüche

1. Verfahren zur Amplifikation einer Ziel-DNA, wobei die DNA zuerst durch die PCR amplifiziert wird, die amplifizierte DNA dann mit einem einzelsträngigen, linearisierten Plasmidvektor mit terminalen Regionen, die den terminalen Regionen der PCR-amplifizierten DNA komplementär sind, in Kontakt gebracht wird, wodurch ein cyclisches Produkt, welches einzelsträngige Sequenzen der Ziel-DNA und den Vektor und zwei doppelsträngige Regionen der überlappenden, terminalen Regionen des Vektors und die PCR-

amplifizierte DNA umfaßt, gebildet wird, und das cyclische Produkt dann in einem Wirtsorganismus eingeschleust wird.

2. Verfahren zur Amplifikation einer Ziel-DNA nach Anspruch 1, wobei die PCR ein zweistufiges Verfahren ist, bei dem ineinander angeordnete Primer verwendet werden, wobei der oder jeder Primer der zweiten Stufe einer Sequenz der Ziel-DNA zwischen den Stellen, die den Primern der ersten Stufe komplementär sind, komplementär ist.

3. Verfahren nach Anspruch 1 oder 2, wobei der oder die Primer der zweiten Stufe einzelsträngige Nucleotidextensionen umfaßt/umfassen, die den terminalen Regionen des Vektors komplementär sind.

4. Verfahren nach einem der vorausgehenden Ansprüche, wobei jede der überlappenden Regionen des cyclischen Produkts mindestens eine Restriktionsspaltstelle umfaßt.

5. Verfahren nach einem der vorausgehenden Ansprüche, wobei ein PCR-Primer Mittel zur Immobilisierung trägt oder bereits immobilisiert ist.

6. Verfahren nach einem der vorausgehenden Ansprüche, wobei die ortsspezifische Mutagenese zwischen der PCR-Amplifikation und der Bildung des cyclischen Produkts durchgeführt wird.

7. Verfahren nach Anspruch 6, wobei die PCR-amplifizierte Ziel-DNA vor der Mutagenese immobilisiert wird.

8. Verfahren nach einem der vorausgehenden Ansprüche einschließlich der anschließenden Stufe der Sequenzierung der Ziel-DNA nach der Clonierung.

9. Verfahren nach einem der vorausgehenden Ansprüche, wobei der einzelsträngige, linearisierte Plasmidvektor durch Immobilisieren des Endes eines Stranges eines doppelsträngigen, linearisierten Vektors unter anschließender Strangtrennung gebildet wird.

10. Verfahren zur Diagnose, wobei die Ziel-DNA für einen physiologischen Zustand charakteristisch ist, und die amplifizierte und clonierte DNA, die nach dem Verfahren nach Anspruch 1 produziert wurde, anschließend charakterisiert wird.

11. Kit zur Amplifikation einer Ziel-DNA, umfassend:
    a. einen linearisierten Vektor in einzelsträngiger oder doppelsträngiger Form oder den doppelsträngigen Vektor in linearisierter Form und an einem Ende eines Strangs davon immobilisiert;
    b. eine Polymerase;
    c. zwei PCR-Primer entsprechend den terminalen Regionen des Vektors; und
    d. Triphosphatnucleoside.

12. Kit nach Anspruch 11, umfassend zusätzlich Mittel zur Sequenzierung der Ziel-DNA einschließlich entweder eines markierten Primers oder markierter Nucleosidtriphosphate.

13. Kit nach Anspruch 11, umfassend zusätzlich Mittel zur ortsspezifischen Mutagenese einschließlich:
    i. eines ortsspezifischen Mutagenese-Primers;
    ii. einer zweiten Polymerase; und
    iii. einer Ligase.

Revendications

1. Procédé d'amplification d'un ADN cible dans lequel ledit ADN est d'abord amplifié par PCR, puis mis en contact avec un vecteur plasmidique linéarisé monocaténaire possédant à ses extrémités des régions complémentaires de celles de l'ADN amplifié par PCR, ce qui aboutit à la formation d'un produit cyclique comprenant des séquences monocaténaires issues dudit ADN cible et dudit vecteur et deux régions bicaténaires issues des régions terminales complémentaires du vecteur et de l'ADN amplifié par PCR; le produit cyclique étant alors introduit dans un organisme hôte.

2. Procédé d'amplification d'un ADN cible selon la revendication 1, dans lequel la PCR comprend deux étapes et utilise des amorces intégrées, l'amorce ou chaque amorce de la seconde étape étant complémentaire d'une séquence de l'ADN cible située entre les sites complémentaires des amorces de la première étape.

3. Procédé d'amplification selon la revendication 1 ou 2 dans lequel l'amorce ou les amorces de la seconde étape comprennent des extensions nucléotidiques monocaténaires qui sont complémentaires des régions terminales du vecteur.

4. Procédé selon l'une des revendications précédentes dans lequel chaque région complémentaire du produit cyclique comprend au moins un site de restriction.

5. Procédé selon l'une des revendications précé-

dentes dans lequel une amorce de PCR porte des moyens permettant de l'immobiliser ou est déjà immobilisée.

6. Procédé selon l'une des revendications précédentes dans lequel la mutagenèse dirigée est effectuée entre l'amplification par PCR et la formation du produit cyclique.

7. Procédé selon la revendication 6 dans lequel l'ADN cible amplifié par PCR est immobilisé avant la mutagenèse.

8. Procédé selon l'une des revendications précédentes incluant une étape ultérieure de séquençage de l'ADN cible après le clonage.

9. Procédé selon l'une des revendications précédentes dans lequel le vecteur plasmidique linéarisé monocaténaire est produit par immobilisation de l'extrémité d'un brin d'un vecteur linéarisé bicaténaire, suivie d'une étape de séparation des brins.

10. Procédé de diagnostic dans lequel ledit ADN cible est caractéristique d'un état physiologique et l'ADN cloné et amplifié produit par le procédé de la revendication 1 est ultérieurement caractérisé.

11. Nécessaire pour l'amplification d'un ADN cible comprenant:

    a) un vecteur linéarisé sous forme monocaténaire ou bicaténaire, ledit vecteur bicaténaire sous forme linéaire immobilisée par une extrémité de l'un de ses brins;
    b) une polymérase;
    c) deux amorces de PCR correspondant aux régions terminales dudit vecteur; et
    d) des nucléosides triphosphate

12. Nécessaire selon la revendication 11 comprenant en outre des moyens de séquençage de l'ADN cible incluant soit une amorce marquée soit des nucléosides triphosphate marqués.

13. Nécessaire selon la revendication 11 comprenant en outre des moyens pour la mutagenèse dirigée

    i. une amorce de mutagenese dirigée
    ii. une seconde polymérase, et
    iii. une ligase.

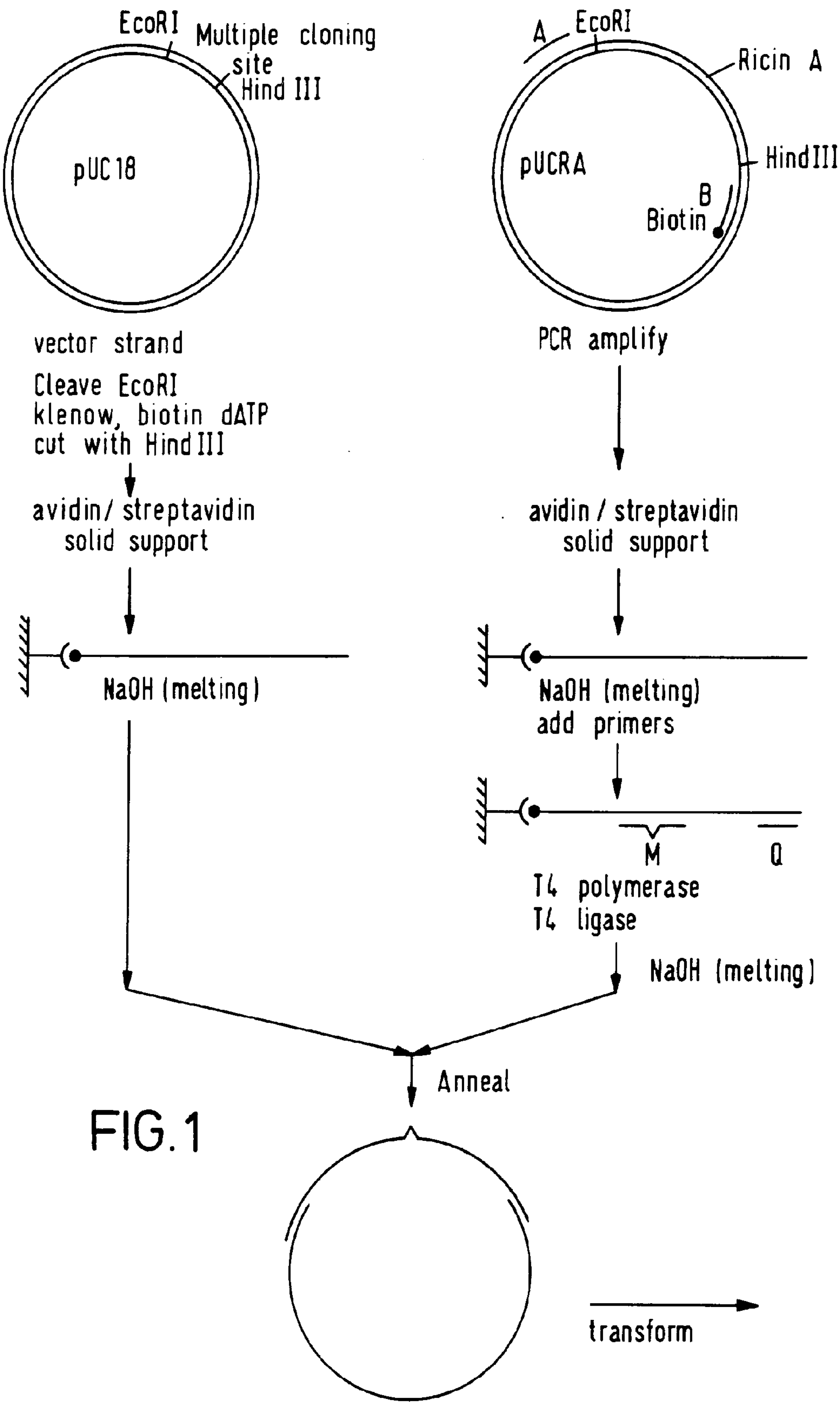
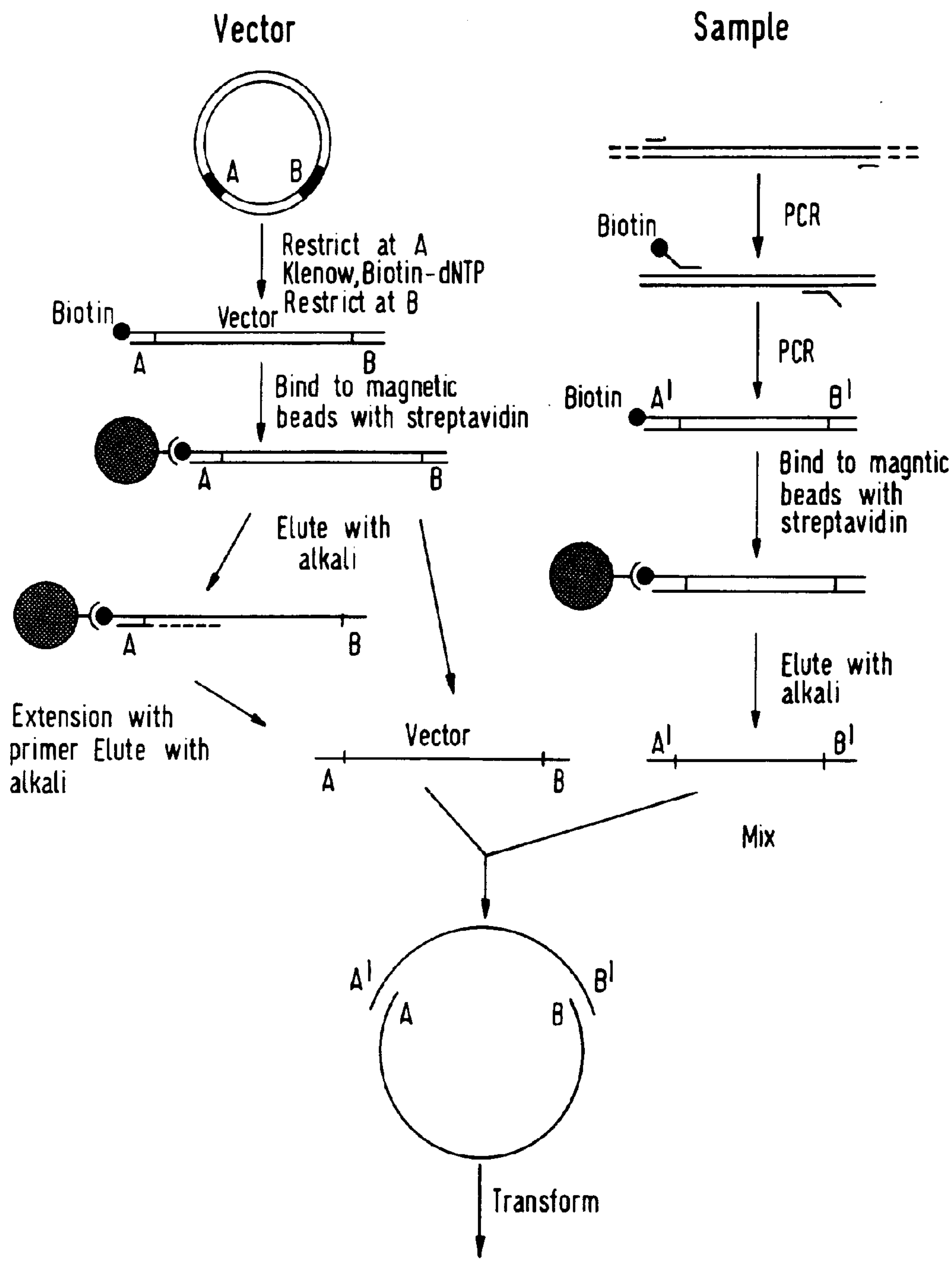

EcoRI
Multiple cloning site
Hind III
pUC 18
A
EcoRI
Ricin A
pUCRA
HindIII
B
Biotin
vector strand
Cleave EcoRI
klenow, biotin dATP
cut with Hind III
PCR amplify
avidin/ streptavidin
solid support
avidin / streptavidin
solid support
NaOH (melting)
NaOH (melting)
add primers
M
Q
T4 polymerase
T4 ligase
NaOH (melting)
Anneal
transform

FIG.1

FIG.2


Vector
Sample
A
B
Restrict at A
Klenow,Biotin-dNTP
Restrict at B
Biotin
Vector
A
B
Bind to magnetic
beads with streptavidin
A
B
Elute with
alkali
A
B
Extension with
primer Elute with
alkali
Vector
A
B
PCR
Biotin
PCR
Biotin
A'
B'
Bind to magntic
beads with
streptavidin
Elute with
alkali
A'
B'
Mix
A'
A
B
B'
Transform

```
3615     RIT113        pUC handle seq 1   RIT125
|
AAGGAGTTGAAGGCCTACAAATCGGAACTGGAGGAACAACTGACCCCGGTGGCGGAGGAGACGCGGGCACGGCTGTCCAAGGAG
LysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGlu

3700                                                        T
|
CTGCAGGCGGCGCAGGCCCCGCTGGGCGCGGACATGGAGGACGTGCGCGGCCGCCTGGTGCAGTACCGCGGCGAGGTGCAGGCC
LeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAla
                                             Cys
                                             112

3784
|
ATGCTCGGCCAGAGCACCGAGGAGCTGCGGGTGCGCCTCGCCTCCCACCTGCGCAAGCTGCGTAAGCGGCTCCTCCGCGATGCC
MetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArgAspAla

3868                                       pUC handle seq 2 Biotin
            T                 RIT123                           RIT114
|
GATGACCTGCAGAAGCGCCTGGCACTGTACCAGGCCGGGGCCCGCGAGGGCGCCGAGCGCGGCCTCAGCGCCATCCGC
AspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArg
            Cys
            158
```

FIG.3

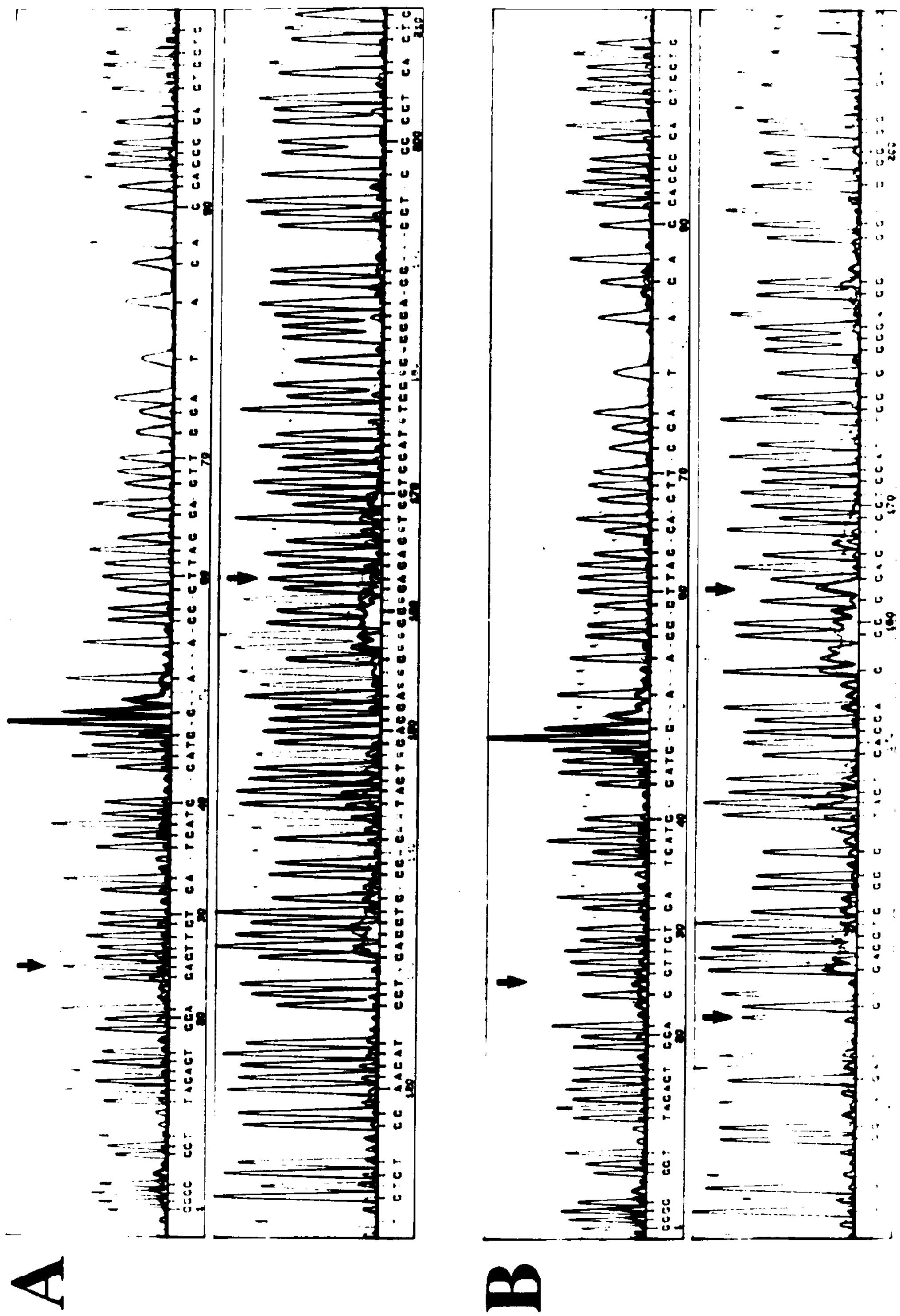


FIG.4